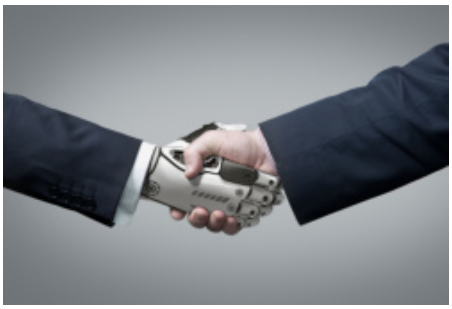
IT時事ネタキーワード「これが気になる！」（第136回）

生成AIガイドライン、「広島AIプロセス」や政府の動向は？

2023.12.28



このコラムで6月に「日本政府、AI戦略に本腰。G7での結果など」という記事を書いた。そこでは、5月に開かれた「G7広島サミット」にて生成AIについての議論が行われ、閣僚級による議論の枠組み「広島AIプロセス」において国際的なルール作りを進めることで各国合意したことを紹介した。

日本政府が世界に先立つ生成AIへのガイドライン。「広島AIプロセス」以降の動向

2023年は生成AIの年、ともいえる。いいも悪いも、生成AIは世の中に大きな変化を与えた。今年も終わろうとしている今、「広島AIプロセス」はどうなっているのか、AIに対する国際的なルール作りはどうなっているのか、状況を見てみよう。

「G7広島サミット」における首脳宣言では、「我々は、信頼できるAIという共通のビジョンと目標を達成するためのアプローチと政策手段が、G7諸国間で異なり得ることを認識しつつも、AIガバナンスに関する国際的な議論とAIガバナンスの枠組み間の相互運用性の重要性を強調する」、そして「我々は、関係閣僚に対し、生成AIに関する議論のために、包摂的な方法で、OECD及びGPAIと協力しつつ、G7の作業部会を通じた、広島AIプロセスを年内に創設するよう指示する」とある。

10月のG7首脳においては「広島AIプロセスに関するG7首脳声明」が発出された。これは、9月の「広島AIプロセス閣僚級会合」にて、G7広島サミットから年末までの中間的な取りまとめとして出された「G7広島AIプロセス G7デジタル・技術閣僚声明」をベースに、「AIのための包摂的なガバナンスの形成をめざした」ものだ。この声明には「広島AIプロセスを更に前進させるための作業計画を年末までに策定する」「11月1日と2日の英国のAI安全性サミットに期待する」とある。

また、声明とともに、AIシステム開発者のための、高度なAIシステムを開発する組織に対する「高度なAIシステムを開発する組織向けの広島プロセス国際指針」や「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」が出された。この国際指針・行動規範では「AIライフサイクル全体にわたるリスクを特定、評価、軽減するため、導入前に適切な措置を講じる」「導入後の脆弱性やインシデントを特定して緩和する」「AIシステムの能力、限界、適切・不適切な使用領域を公表、十分な透明性の確保を支援する」「電子透かしなど、ユーザーがAIが生成したコンテンツを識別できるようにする」「個人データおよび知的財産を保護する」など、11の原則（両者とも項目は同じ）が示されている。

そして、前述した英国での「AI安全性サミット」には、岸田首相がオンライン形式で参加し、「広島AIプロセス」において、国際的なルール作りに取り組んでおり、グローバルなAIのルールの共通の基盤となると確信していると述べた。さらに、10月の首脳声明および国際指針、国際行動規範を紹介しつつ、年末にかけて「広島AIプロセス包括的政策枠組」の策定に向けた作業を加速させるとともに、「広島AIプロセス」は「AI安全性サミット」の取り組みとも相互補完的であるとし、引き続き密に連携していきたい、としている。

G7議長国として主導してきた「広島AIプロセス」。次のG7議長国イタリアに引き継ぎ

続く11月に開かれた第6回「AI戦略会議」では、年末に向けての「広島AIプロセス」の作業の加速化と「広島AIプロセス包括

的政策枠組」策定の加速化、およびG7以外も含んだマルチステークホルダーとの協議の実施など、「広島AIプロセス」をさらに前進させるための作業計画の策定を担当閣僚に求めている。

そして12月1日。政府は「G7デジタル・技術大臣会合」をオンラインで開催し、「広島AIプロセス G7デジタル・技術閣僚声明」を採択した。声明では、日本のG7議長国下での「広島AIプロセス」の集大成として「広島AIプロセス包括的政策枠組み」を承認した他、「広島AIプロセスを前進させるための作業計画」「全てのAI関係者向けの広島プロセス国際指針」などを取りまとめている。

作業計画では「我々は、来年のイタリアのG7議長国の下、AIに関する取組を継続する予定」とし、「関連国の政策動向及び国際行動規範にコミットする組織のリストに関する最新情報を提供するために、広島AIプロセスの専用ウェブサイトを立ち上げる予定」など、イタリア議長国の下で引き続き協力していく姿勢を明らかにした。

今回の「全てのAI関係者向けの広島プロセス国際指針」では、これまでの11の原則に「高度なAIシステムの信頼でき責任ある利用を促進し、貢献する」を加え、「AI関係者は、高度なAIシステムが特定のリスクをどのように増大させるか及び／又は新たなリスクをどのように生み出すかといった課題を含め、自分自身そして必要に応じて他者のデジタル・リテラシー、訓練及び認識を向上させる機会を求める」「高度なAIシステムの新たなリスクや脆弱性を特定し、それに対処するために、必要に応じて、協力し情報を共有する」ことを推奨している。

12月6日には「G7首脳テレビ会議」がオンラインで開かれ、G7日本議長年の総括としてウクライナ情勢や中東情勢をはじめとする重要課題について議論が行われた。AIについては12月1日の「G7デジタル・技術大臣会合」で合意した「広島AIプロセス包括的政策枠組」について、「世界で初めてAI関係者が遵守すべきルールを包括的に定めた画期的なもの」と強調。続いて首相は、広島AIプロセスでの具体的な成果により、急速に進展する生成AIに対して、G7が効果的かつ迅速に対応できることを世界に力強く示すことができたこと、これで広島AIプロセスは終わりではなく、今後の作業計画に基づき、国際指針や行動規範など今般の成果を広く国際社会に拡大していくこと、来年のイタリア議長下でも信頼できるAIの実現に向けた取り組みを引き続きG7で主導していくこと、などを述べている。

「広島AIプロセス」以外のAIへの動き。世界に対する日本の位置など

日本では「広島AIプロセス」の他、首相官邸における「知的財産戦略本部」での「AI時代の知的財産権検討会」や、経済産業省における「AI事業者ガイドライン検討会」「AIガバナンス」、内閣府における「AI戦略チーム（関係省庁連携）」「人間中心のAI社会原則会議」、デジタル庁における「デジタル社会推進会議」など、AIに関するさまざまな活動が行われている。

生成AIが業務でも普通に活用されつつある今、民間の動きもある。2023年10月26日、生成AIを中心とするAIのビジネス利用が急拡大する中でのリスク認識の高まりと、それを受けた政策的な動向を踏まえ、東京海上やリクルートなど20社以上の有志企業により「AIガバナンス協会（AIGA）」が設立された。協会では、企業と社会が安心してAIを活用し、持続可能な成長を遂げるために、多様なプレーヤーがAIガバナンスのあり方を議論するとともに、そうした知見をもとに政策提言等の活動を行うという。

諸外国の動きとしては、EU（欧州連合）の主要機関が12月9日、2021年4月から提案されているAI規制法が合意に至ったと発表し、AIを規制するための世界初の法案と称している。AI規則案は「規則」（Regulation）とし、すべての加盟国に統一ルールが直接適用される。2024年後半に完全施行をめざすとしている。

このAI規制法では、AI使用によるリスクを「許容できないリスク」「ハイリスク」「限定リスク」「最小リスク」の4段階に分けたリスクベースアプローチを採用し、個別の制約を設ける。生成AIについては、生成物に関してAIによる生成を明記することや元データを開示するなど、透明性の義務が課される。対応を怠った企業には、巨額の制裁金などが科される。また「AIオフィス」と呼ぶ監督機関を設け、厳格な法律適用を図る。いわゆる「ハードロー（議会などで成立した法律や条例、法的義務のある条約などの法令で、違反すれば刑事罰や行政処分などの法的制裁を受ける）」での管理といえる。

また中国においては、8月15日に、中国初の生成AIに対する法令である「生成人工知能サービス管理暫定弁法」が施行され、生成AIに対しての基本規範を確立したという。すでに制定されている「中国サイバーセキュリティ法」「中国データセキュリティ法」「中国個人情報保護法」「中国科学技術進歩法」などに準拠、法というよりはガイドラインに近い内容という。

英国では先に触れたとおり、11月に「AI安全性サミット」が行われた。AI安全性サミットは、AIの急速な発展を踏まえ、AI技術の安全な開発と使用に関する会合として英国が立ち上げたものだ。各国政府や、関係国際機関、民間企業、研究者などが参加（日本からは小森総務大臣政務官が出席）、AIにおける安全性などについて議論を行った。サミットには前述のご

とくオンラインで岸田首相がコメントを寄せるなど、広島AIプロセスと密に連携する方向だ。

AI研究開発をけん引するアメリカは、AI規制について、基本的に「ソフトロー（法的拘束力を持たない、または弱い法律）」による対応を進めており、ガイドラインなどによるゆるやかな規制への誘導という戦略を採っている。米国科学技術政策局は、AIを含む自動化システムの設計、使用、導入の指針となるべき「AI権利章典」を昨年10月に公表し、「安全で効果的なシステム」「アルゴリズム由来の差別からの保護」「データのプライバシー」「ユーザーへの通知と説明」「人による代替手段、配慮、フォールバック」という5つの原則を掲げる。

その他、今年1月、米国立標準技術研究所（NIST）により、AI技術のリスク管理のためのガイダンス「AIリスクマネジメントフレームワーク（AI RMF）」が発表されている。その前半では「AIに関わるリスクの考え方」や「信頼できるAIシステムの特徴」、後半では「AIシステムのリスクに対処するための実務」が解説され、生成AIに限らず、一般的なAIのリスク管理手法として注目されている。

今後どうなる、傾向と対策

AI統制について、あくまでガイドライン、規範、指針などを示すソフトローで対応する日本や英国などに対し、法的拘束力と罰則の厳しいハードローで対応するEUといえるが、方向がはっきり分かれてきている。ただし、アメリカはハードローとソフトローを両方用いる傾向が見えてきている（一般的にはソフトローな姿勢だが、最近はハードロー的な動きも見られるという）。

考えるに、ソフトローのみで対応できれば言うことはないが、世の中なかなかそうはいかない。ガイドラインだけでは、悪用や犯罪を防ぐことはできないからだ。12月8～10日の「G7茨城水戸内務・安全担当大臣会合」では、G7として初めて組織詐欺を議題に取り上げ、国境を超えた特殊詐欺などの組織犯罪や生成AIを使った犯罪に対して、各国で協力して取り組んでいくことで一致したという。犯罪を防ぐには、ハードローも必要と思う。

最近は、フィッシング詐欺などにAIで生成された文章や画像が使われることが多く、手口が高度化、被害も拡大傾向とされる。AIで作成した写真やフェイク動画などは見分けをつけることは難しい。AIが犯罪に使われる可能性を考えると、ガイドラインや規範、原則、電子透かし、などでは足りない気もする。何らかの法的拘束力と罰則のあるハードロー的対策も今後は必要になっていくだろう。

国際的な業務を行う場合には、各国の法が適用される場合もあり、注意が必要だ。EUの規制法が成立すれば、EUに住む人を対象としたサービスであれば、日本の業者も制裁金などが課される可能性も考えられる。最近は仕事にAIを使った場合はその旨を記す、などが契約に含まれる場合も多い。さまざまな変化への対応が必要だろう。

日常生活においても、AI利用による犯罪に巻き込まれる、偽情報に惑わされる、見分けがつきにくい情報に困惑するなどの事態は誰しも遭遇する可能性がある。情報を定期的に収集して、見守っていくしかない。広島AIプロセスその他、各国の動向に注目していこう。

※掲載している情報は、記事執筆時点のものです